pesquisa fapesp

N.359__JAN 2026__ANO 27

O desafio de ganhar o mundo

Startups brasileiras que buscam atuação internacional enfrentam barreiras culturais, tecnológicas e regulatórias

Vacina do Butantan contra a dengue chega aos profissionais da saúde

Cientistas e criadores se mobilizam para salvar os jumentos da extinção no país

Novos fósseis revelam variedade de tamanho e de hábitos de pterossauros

Embrapa obtém autorização especial para pesquisa com cultivo de cannabis

Com maior visibilidade, arte indígena requer formas alternativas de curadoria

Fronteiras da avaliação

Atlas britânico mapeia os sistemas que medem a qualidade da pesquisa de diversos países

SARAH SCHMIDT __ ilustrações **MARIANA ZANETTI**

Como os países medem o desempenho de cientistas e instituições de pesquisa? O que seus sistemas de avaliação têm em comum? Essas são algumas perguntas que o observatório on-line Atlas da Avaliação, do Research on Research Institute (RoRI), sediado no Reino Unido, procura responder. A plataforma, lançada em maio de 2025, já compilou informações sobre os modelos nacionais de avaliação de 17 países e mostra como eles evoluíram nos últimos anos. Filtros permitem selecionar os modelos adotados de acordo com diferentes características. Se o interesse do usuário for identificar sistemas que aferem a qualidade e a reputação de instituições de pesquisa a fim de orientar a aplicação de recursos, o Atlas destaca dez países – México, Portugal, Itália, China, Austrália, Polônia, República Tcheca, Letônia, Reino Unido e Brasil. Em outras duas nações, Índia e Colômbia, os modelos de avaliação têm um objetivo mais restrito: produzir estatísticas e oferecer uma visão geral da atividade de pesquisa. Caso a busca seja por sistemas que fazem uma avaliação individual de pesquisadores, são apontados quatro países: México, Colômbia, Argentina e China.

A plataforma também permite ver em profundidade o modelo de cada nação. Alguns países têm mais de um, como a China, que utiliza quatro sistemas que se complementam. A Avaliação Nacional de Disciplinas da China analisa o desempenho do conjunto de pesquisadores de cada campo disciplinar em ciclos de quatro a cinco anos – a participação das instituições é voluntária. Já o Plano Double First-Class, lançado em 2015 pelo governo chinês com a ambição de estabelecer universidades e disciplinas de padrão mundial até 2050, tem uma influência forte no financiamento de instituições de elite. Há um sistema que avalia e apoia pesquisadores com produção científica destacada, oferecendo salários elevados, financiamento e suporte institucional, e outro que mede o desempenho de mais de 100 institutos vinculados à Academia Chinesa de Ciências e é usado para distribuir recursos.

A Polônia é outro país que combina vários sistemas. Dois deles têm forte influência na concessão de financiamento: um se concentra na qualidade da pesquisa das universidades e outro, em seu desempenho institucional, analisando o número de estudantes atendidos e a existência de programas de internacionalização. Uma terceira iniciativa avalia as universidades com melhor desempenho em indicadores de pesquisa – as 10 mais destacadas recebem reforço nos investimentos.

A maioria dos países mapeados no Atlas mantém sistemas de avaliação unificados. O Reino Unido dispõe do Research Excellence Framework (Ref), que analisa o impacto e a qualidade da pesquisa produzida pelas instituições de ensino superior britânicas e o ambiente de trabalho que elas oferecem. Apesar de voluntária, a participação das instituições é praticamente universal, devido ao seu forte impacto na reputação e no financiamento – anualmente, mais de £ 2 bilhões são distribuídos com base em seus resultados. Cada pesquisador pode submeter até quatro trabalhos para avaliação, entre os mais relevantes que produziu em determinado período. O último exercício de avaliação foi divulgado em 2022 e o próximo virá

em 2029. O modelo britânico serviu de inspiração para países como Portugal, Itália, Noruega e Eslováquia.

No Brasil, desde 1976 a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) promove a avaliação dos programas de pós-graduação brasileiros para aferir a qualidade dos cursos de mestrado e doutorado, acadêmicos e profissionais, e nortear a distribuição de bolsas e recursos para programas. Os ciclos de avaliação são de quatro anos. No mais recente deles, divulgado no fim de 2022, 4.512 programas *stricto sensu* submeteram para avaliação 4,7 milhões de produções intelectuais de seus pesquisadores entre 2017 e 2020 (*ver* Pesquisa FAPESP *nº 324*). A publicação dos resultados da avaliação quadrienal 2021-2024 está prevista para ocorrer em janeiro de 2026. Os programas recebem notas que vão de 3 a 7. Enquanto 3 é o valor mínimo para que o programa continue existindo, notas entre 6 e 7 indicam excelência e nível de competitividade internacional. O Brasil também conta com a modalidade de bolsas de produtividade em pesquisa do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), que classifica os pesquisadores de maneira individual para concessão de apoio – em 2025, havia 17.080 beneficiários. Há ainda mecanismos no âmbito estadual, criados pelas fundações de amparo à pesquisa.

O modelo do México se debruça sobre o desempenho individual de cientistas. Seu Sistema Nacional de Pesquisadores (SNI), coordenado pelo Conselho Nacional de Ciência e Tecnologia do país, ocorre em ciclos de cinco a 10 anos, dependendo do nível do cientista, e impacta diretamente a progressão na carreira e acesso a financiamento público. "Ao definir os critérios das avaliações, os sistemas influenciam o comportamento dos pesquisadores e suas organizações e, de certa forma, moldam o que vai acontecer no mundo da ciência", observa Sergio Salles-Filho, coordenador do Laboratório de Estudos sobre Organização da Pesquisa e da Inovação da Universidade Estadual de Campinas (Unicamp). Salles-Filho é um dos 26 autores de um relatório publicado em maio de 2025, que analisou os sistemas de 13 países, apresentou a metodologia desenvolvida pelo RoRI para distinguir os tipos de modelos, e serviu de base para organizar o Atlas.

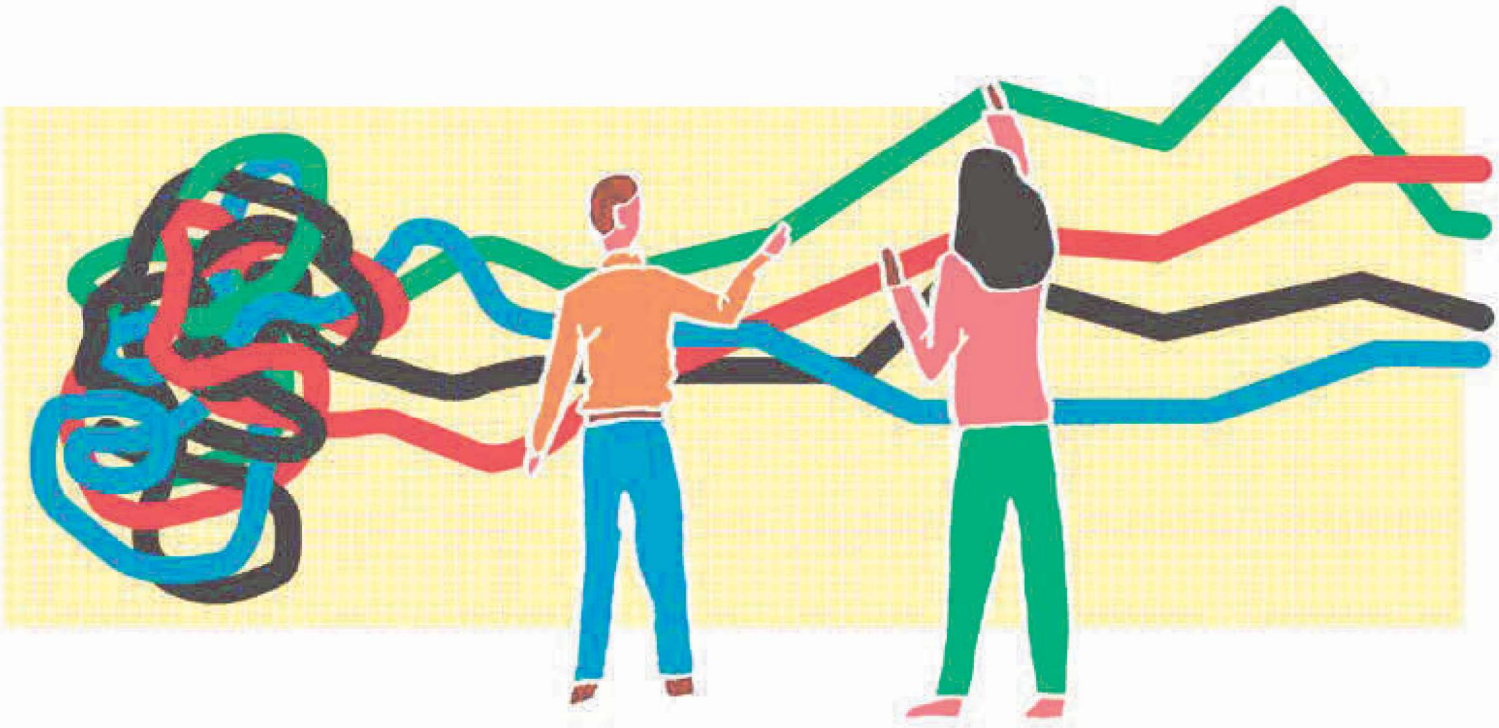
AO LONGO DO TEMPO

Uma das conclusões do estudo é que não existe uma receita de avaliação a ser seguida, ainda que os países possam aprender com a experiência dos demais. O trabalho também mostra como os sistemas evoluíram nos últimos 40 anos e aponta três grandes momentos históricos. Em boa parte do século XX, o padrão, ditado pelo Reino Unido e grande parte da Europa, era avaliar a pesquisa de uma forma compartimentada, com foco em cada campo disciplinar. Nos anos 1990, consolidou-se um novo perfil, que analisa o desempenho dos sistemas nacionais com base em indicadores de produtividade e critérios de excelência, estimulando a competição por recursos. Esse ainda é o modelo dominante. "Hoje, por conta de distorções nos sistemas de avaliação, discute-se a emergência de um novo paradigma, o da avaliação responsável", destaca o pesquisador da Unicamp. O mote dessa nova fase é adotar critérios de avaliação mais abrangentes, levando em conta também dimensões como o impacto econômico e social da pesquisa, para além dos indicadores de produtividade científica.

Salles-Filho explica que esse tema começou a ganhar força nos últimos anos, na esteira da Declaração de São Francisco sobre Avaliação da Pesquisa (Dora), publicada em 2013, que defende a superação do uso do fator de impacto (FI) e de métricas de periódicos como medidas de qualidade da pesquisa. Calculado pelo número médio de citações recebidas pelos artigos de um periódico, o FI é adotado para examinar o prestígio de revistas científicas. Em 2015, veio o Manifesto de Leiden, reforçando que métricas devem apoiar e não substituir a avaliação qualitativa. Mais recentemente, em 2022, surgiu a Coalition for Advancing Research Assessment (Coara), coalizão global de instituições comprometidas com a reforma da avaliação da pesquisa, também em direção a modelos mais amplos e diversificados.

Um sistema de avaliação que se destacou na crítica aos indicadores de produtividade é o dos Países Baixos. O protocolo de avaliação das universidades adotado no país em 2021, e que está vigente até 2027, proíbe o uso de indicadores como o FI. Já o índice-h, que combina o número de artigos de um autor com a frequência de citações desses manuscritos, precisa de justificativa para ser considerado. Apesar da ousadia, a influência do modelo é restrita. "Nos Países Baixos, praticamente não há impacto direto da avaliação no financiamento de grupos de pesquisa ou departamentos de universidades", explica André Brasil, pesquisador do Centro de Estudos de Ciência e Tecnologia da Universidade de Leiden, na Holanda. Servidor da Diretoria de Avaliação da Capes, ele é um dos autores de um artigo publicado em abril na revista *Research Evaluation*, que comparou os sistemas de avaliação nacionais do Brasil e dos Países Baixos.

Segundo ele, o foco principal do modelo holandês, que ocorre a cada seis anos, é analisar as



estratégias das unidades de pesquisa e departamentos de universidades e sua capacidade de alcançar seus objetivos, tendo como base uma autoavaliação em que refletem sobre suas características. Pode haver, inclusive, resultados negativos, como constatar que um plano promissor não funcionou, o que, destaca o pesquisador, é parte da ciência, já que o "não resultado" também é um resultado. Há ainda o trabalho de uma comissão externa, internacional, que dá orientações às unidades para melhorar sua atuação e atingir objetivos. "É um modelo interessante. Funciona bem em um país pequeno, com um sistema integrado e um nível de investimento em pesquisa proporcionalmente mais alto do que no Brasil. Como a avaliação interfere pouco na distribuição de recursos geralmente escassos, ela se torna um exercício de aprendizagem e reflexão", explica.

"A autoavaliação passou a ser considerada como componente das últimas avaliações da Capes, mas ainda com peso reduzido em relação ao seu potencial transformador", complementa. No próximo ciclo de análise dos programas de pós-graduação brasileiros, haverá novidades importantes. O Qualis-periódicos, sistema de classificação de revistas científicas criado nos anos 2000, não será mais usado para avaliar a produção de estudantes e de docentes dos programas de pós-graduação. A análise passa a se concentrar, também, na qualidade e na classificação do artigo e não somente no desempenho da revista científica em que ele foi publicado (*ver* Pesquisa FAPESP *nº 345*). Outra mudança é a introdução dos chamados "casos de impacto". Será possível apresentar aos avaliadores relatos de como as pesquisas geraram impacto científico, educacional, social, econômico e em políticas públicas.

Na avaliação de Jacques Marcovitch, reitor da Universidade de São Paulo (USP) entre 1997 e 2001, expedientes como o uso de casos de impacto representam uma mudança de cultura no sentido de aproximar da sociedade o conhecimento produzido pelos programas de pós-graduação. "Hoje, só se pede a um doutorando ou mestrando para realizar sua pesquisa e publicá-la. Mas será que essa produção alcança ministérios, prefeituras, empresas ou ONG?", pergunta. "Se o estudante fizer uma síntese de quatro páginas, narrando o efeito que seu trabalho pode produzir na sociedade, é mais provável que seja levado em conta por tomadores de decisão."

Marcovitch lidera o Projeto Métricas, apoiado pela FAPESP, iniciativa que reúne pesquisadores e gestores de diversas instituições para desenvolver formas abrangentes de monitorar o impacto das universidades na sociedade. A abordagem comparativa propiciada pelo Atlas é bem-vinda, na sua avaliação. "Temos que aprender uns com os outros, mas a plataforma ficará melhor quando tiver mais países incluídos", observa. Ele vê espaço para aprimoramento nos filtros do Atlas. Segundo observou, eles mostram dimensões de como a ciência é avaliada, mas poderiam analisar melhor como a sociedade se apropria dela. ●

Os artigos científicos consultados para esta reportagem estão listados na versão on-line.